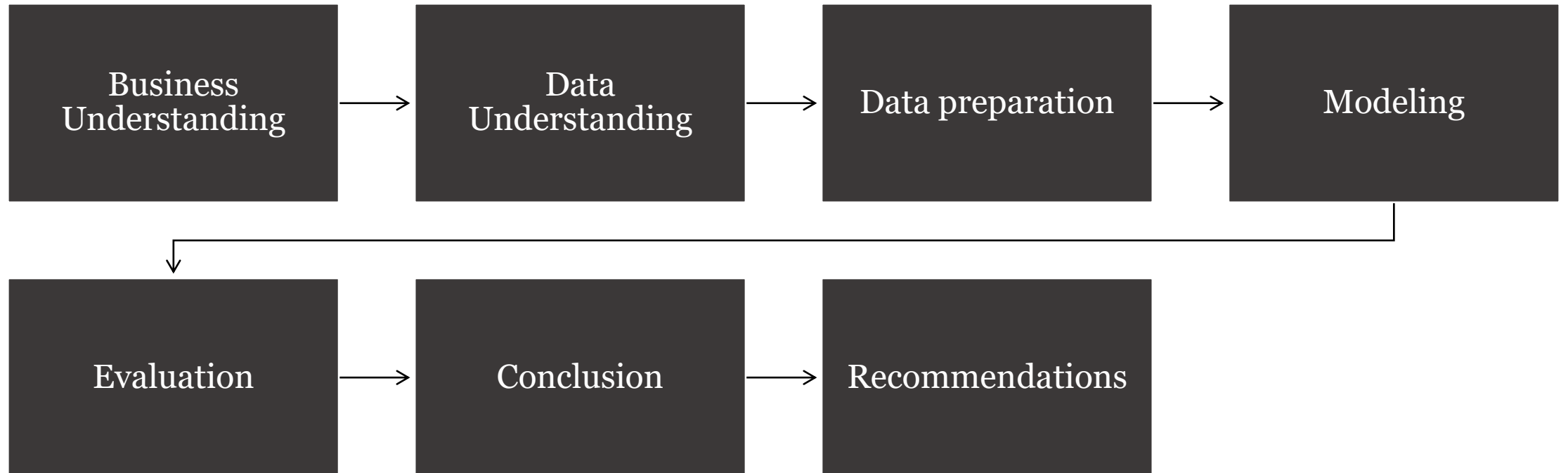


# CUSTOMER CHURN PREDICTION



# Outline



# Overview



- Customer churn is a critical focal point in the telecommunications industry
- It costs **5 times more** to acquire a new customer than to keep an existing one
- For telecoms, gaining a firm understanding of churn rates and prioritizing customer retention sets the stage for sustained profitability and long-term success

# Business Problem

- Syriatel Telecommunications is experiencing a substantial increase in customer churn rates and wants to understand the underlying factors contributing to this trend



# Objectives



To understand which factors or variables contribute the most to customer churn



To identify different customer segments based on churn behaviour



To develop a model that can accurately predict customer churn



To obtain valuable insights that help generate the best recommendations to protect Syriatel's revenue

# Data understanding

- The analysis investigates a telco dataset containing **3333** records and **21** variables
- The features encompass a mix of numerical and categorical variables
- The target variable is "**churn**"

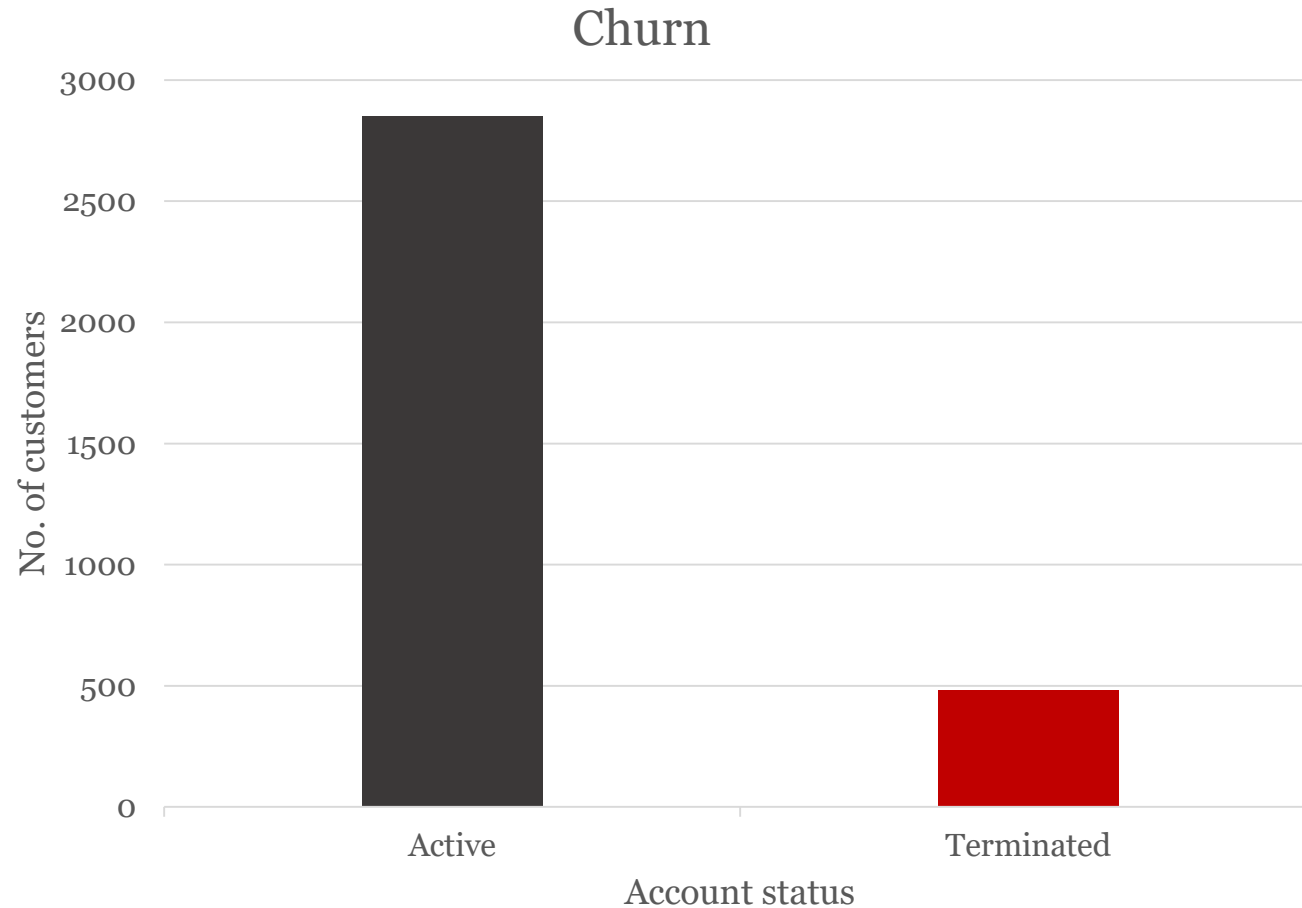


# Data preparation

- The dataset underwent cleaning, including checking for duplicated rows and missing values, and dropping unnecessary columns and
- Data analysis included univariate analysis of the target variable "**churn**" and bivariate analysis of churn in relation to other features in the dataset



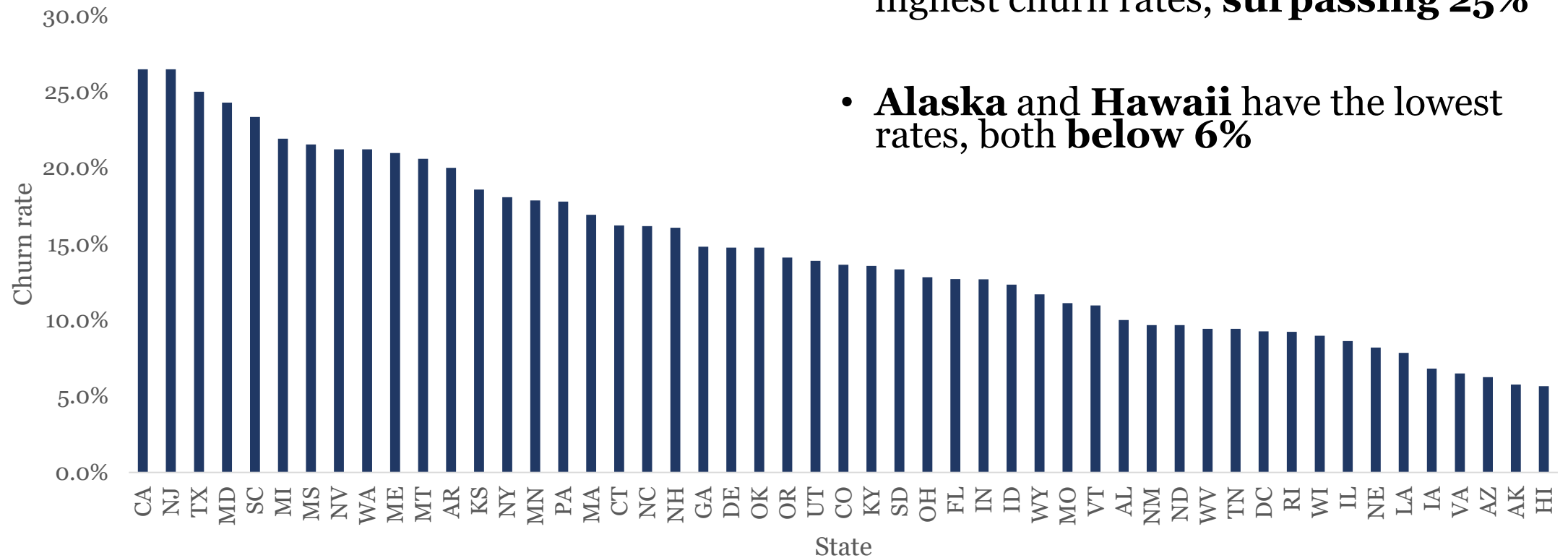
# Churn vs. Retention



- Total number of customers: **3,333**
- Terminated accounts: **483**
- Churn rate: **14.5%**



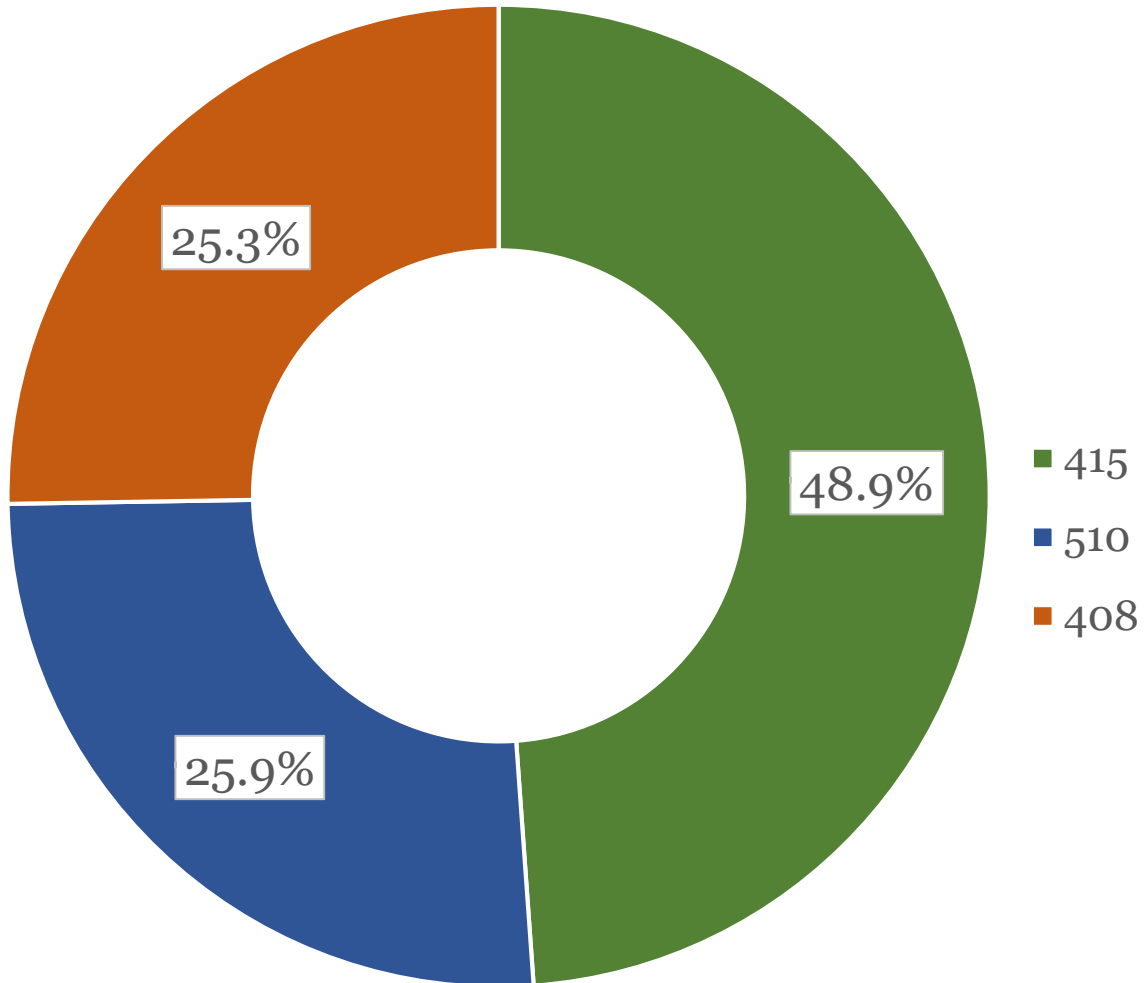
# Churn by State



- **California** and **New Jersey** exhibit the highest churn rates, **surpassing 25%**

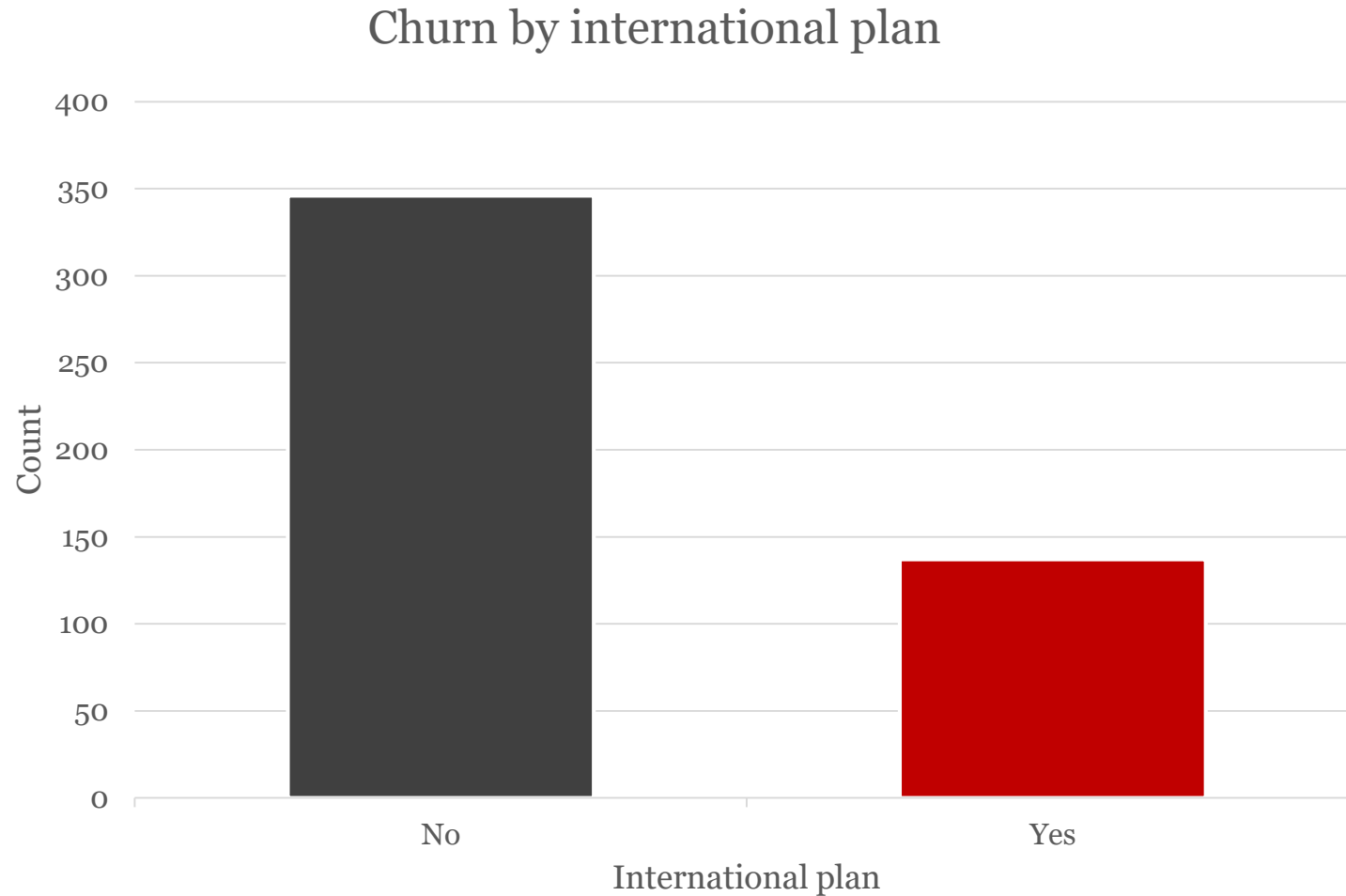
- **Alaska** and **Hawaii** have the lowest rates, both **below 6%**

# Churn by area code



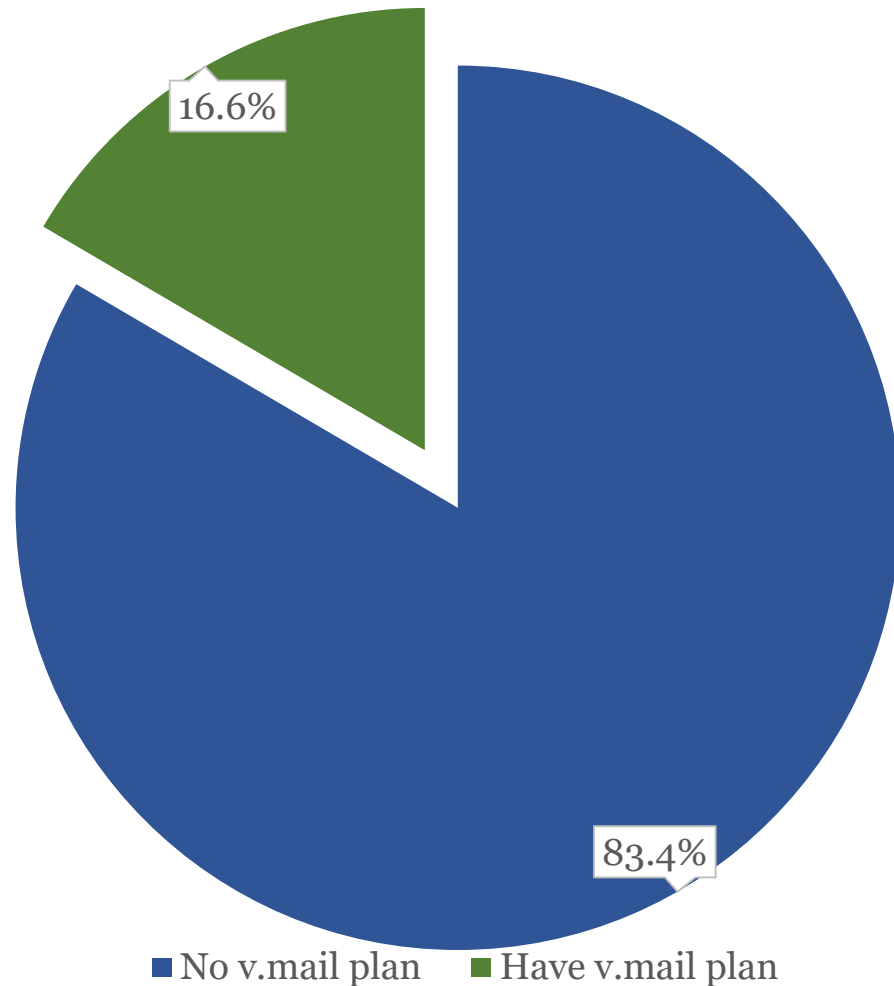
**Area code 415** exhibits the highest percentage of account terminations

# How does an international plan affect churn?



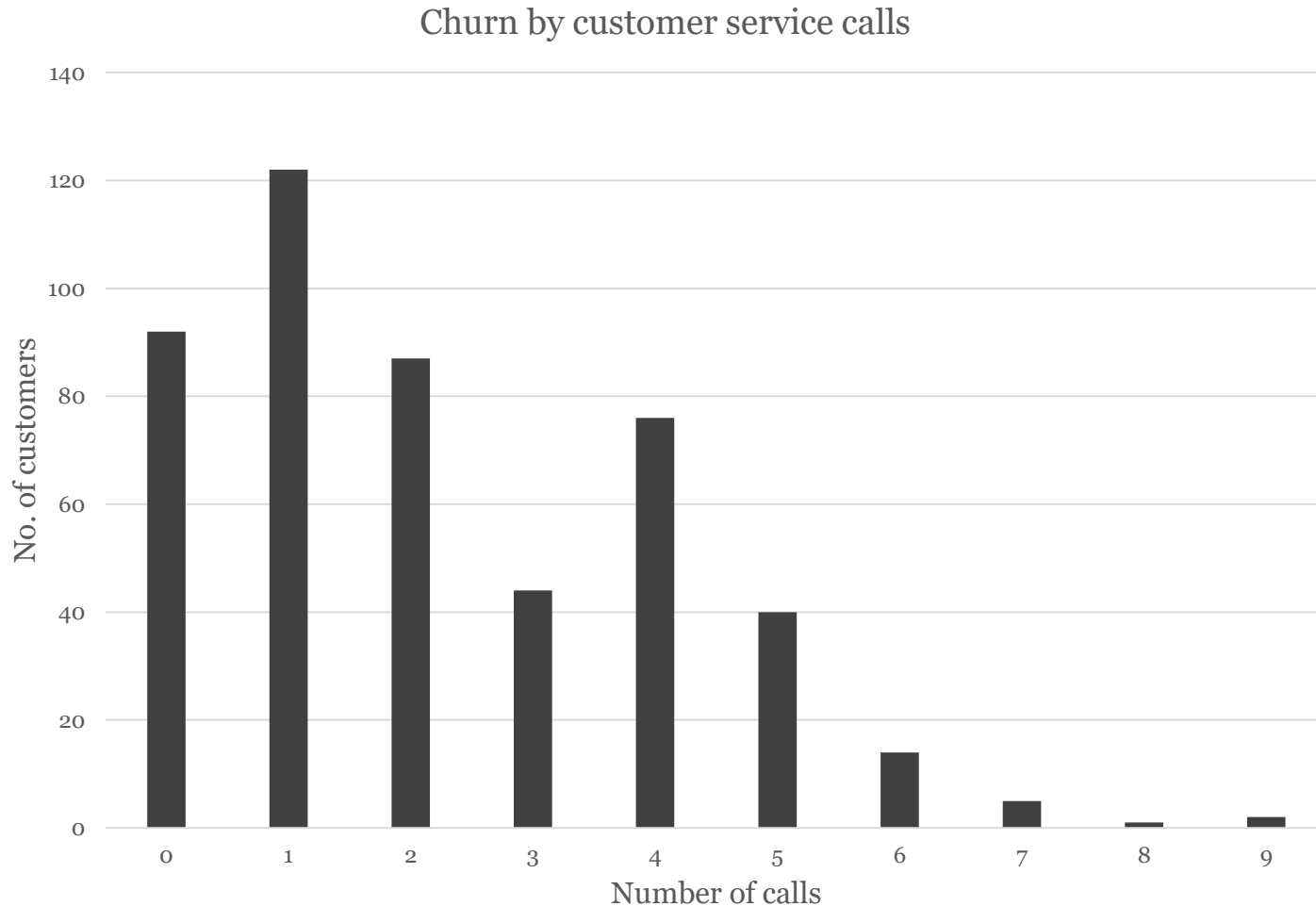
- Of the 483 customers that terminated their accounts, **346** did not have an international plan

# How does a voicemail plan affect churn?



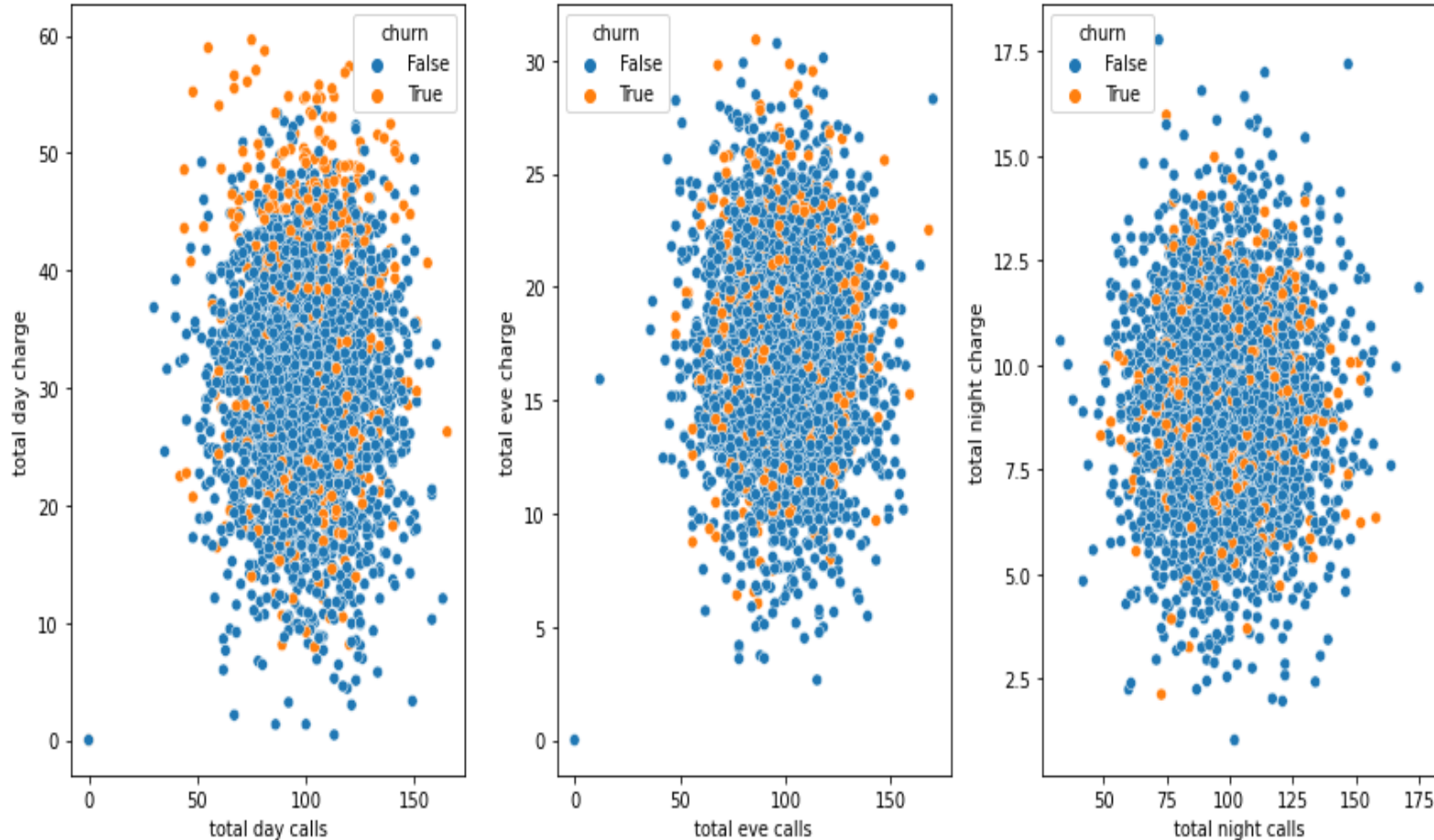
- Of the 483 customers that terminated their accounts, **83.4%** did not have a voicemail plan

# Churn by customer service calls



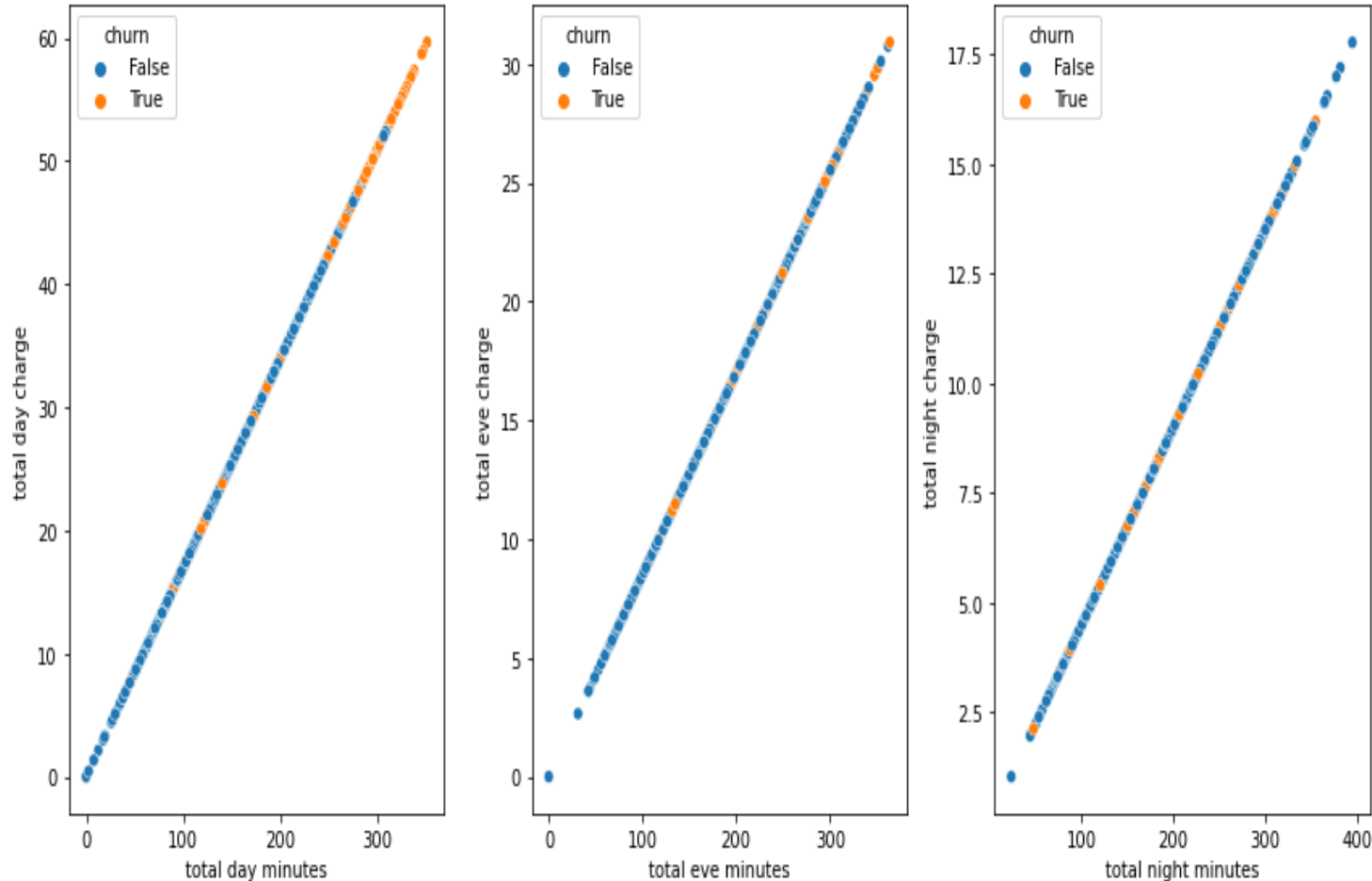
- Majority of account terminations are associated with customers who made only one call to customer service
- A considerable proportion of customers terminated their accounts without contacting customer service

# Churn by Total calls vs. Total charges



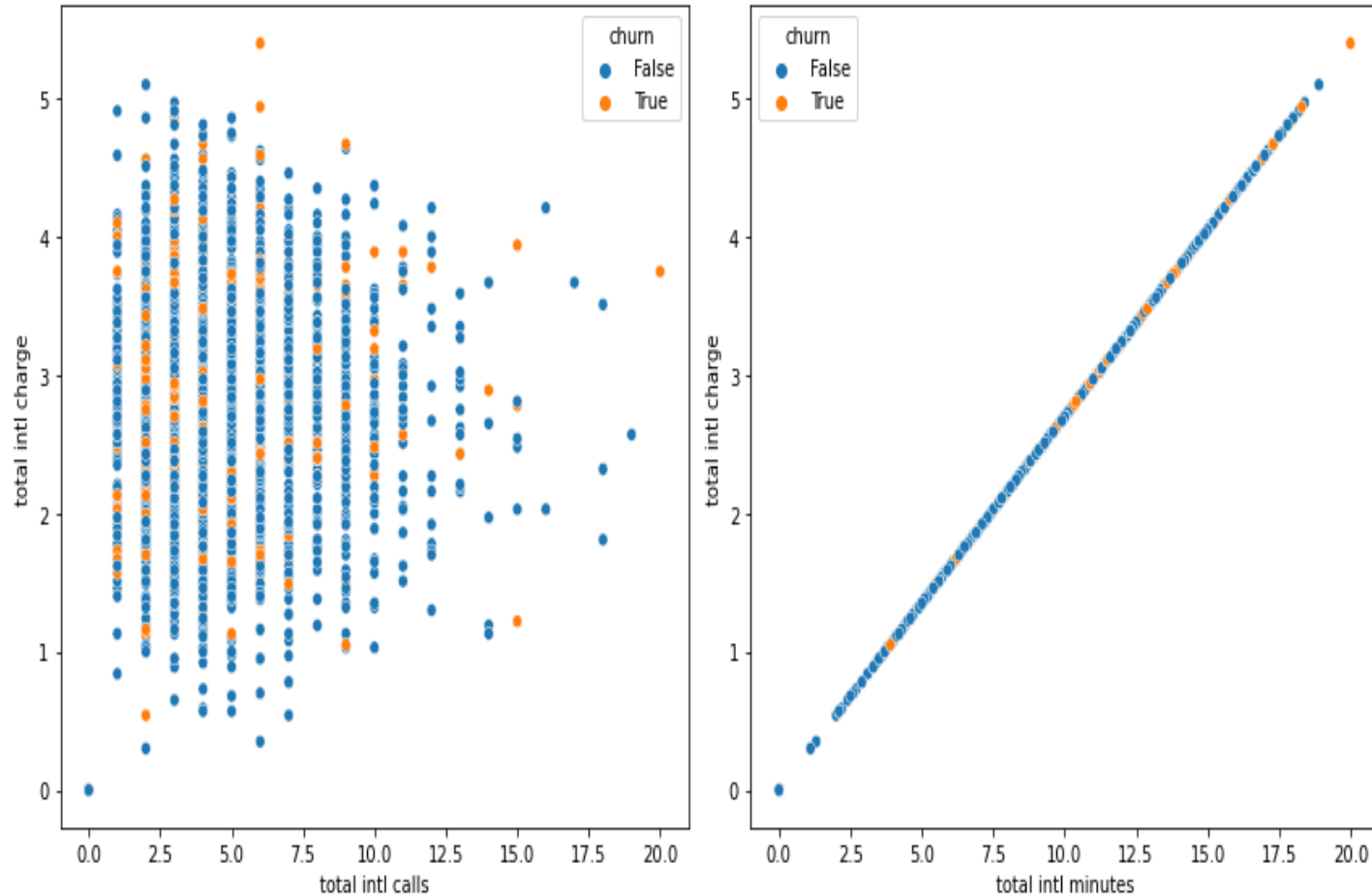
- Churn rates are higher during the day due to the relatively higher call charge rates compared to the cheaper evening and night calls
- Call charges for daytime, evening, and nighttime are higher even with fewer calls made

# Churn by Total minutes vs. Total charges



- On average, customers who terminated their accounts appear to have subscribed to more day minutes, leading to higher charges

# Churn by Total international calls and minutes vs. Total international charges



- A significant number of account terminations are linked to higher charges on fewer international calls



# Modeling

Modeling algorithms used:

Logistic Regression

Decision Tree

Random Forest

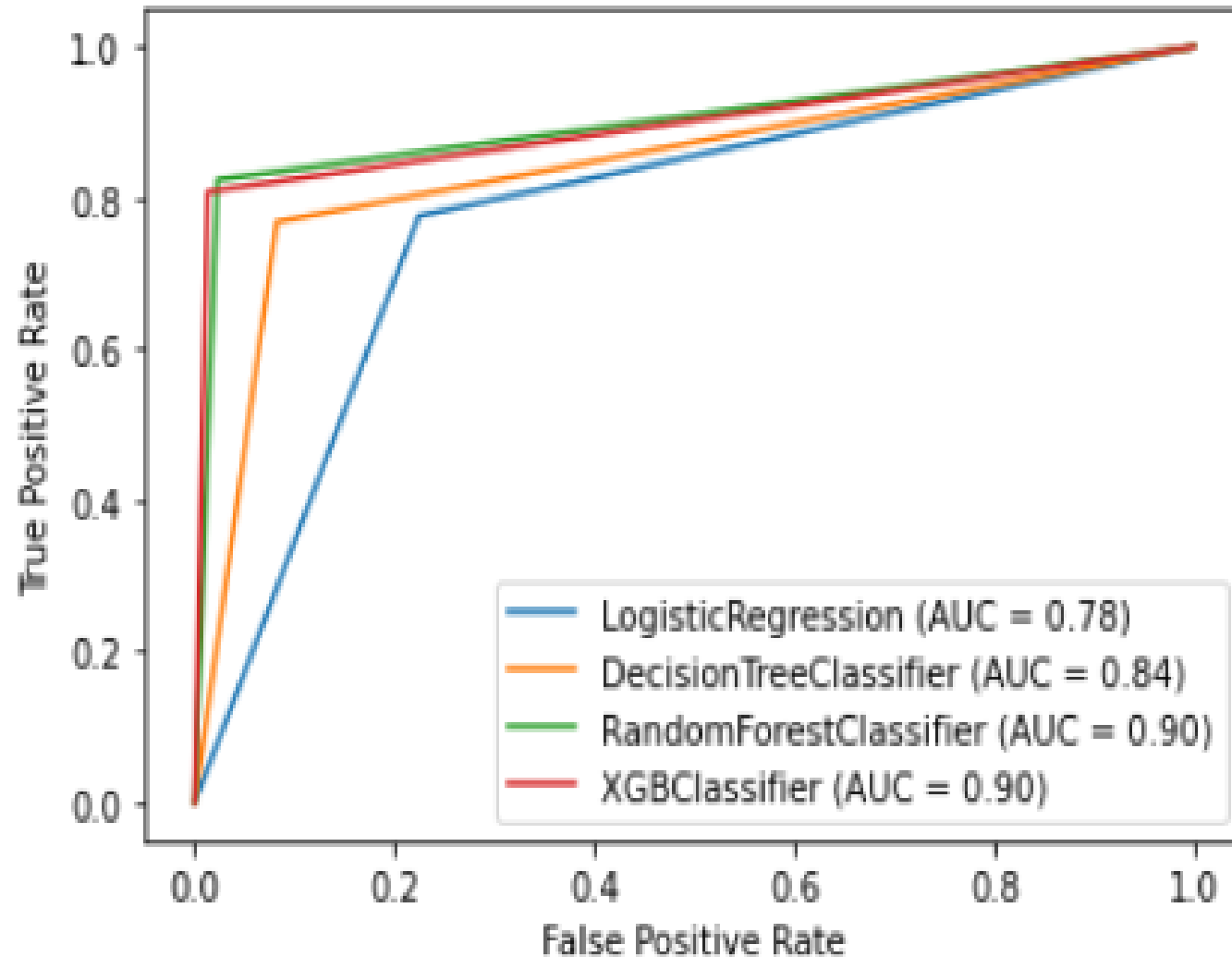
XG Boost

# Evaluation

Based on:

- **ROC\_AUC curve** – shows how efficient the model is in distinguishing between customers who churned and those who did not
- **Recall** - quantifies the model's capacity to accurately detect customers who are likely to churn among all the customers who churn
- **Accuracy** – measures the ability of the model to classify customers correctly

# ROC-AUC results



Out of the 4 algorithms used, **XGB classifier** and **Random Forest Classifier** achieve the highest efficiency rates in churn prediction at **90%**

# Recall results

	accuracy	recall
classifiers		
LogisticRegression	0.775779	0.776
RandomForestClassifier	0.954436	0.832
DecisionTreeClassifier	0.894484	0.744
XGBClassifier	0.960432	0.808

Based on recall,  
Random Forest  
Classifier achieves  
**83.2%** accuracy in  
detecting customers  
that are likely to churn

# Winning model: Tuned Random Forest Classifier

1

Surpassed other models with an **89% AUC score**, highlighting its robust capability to distinguish between churned and active customers

2

Has a good balance between **sensitivity** (recall) and **specificity**, capturing a high proportion of both churned and active customers accurately

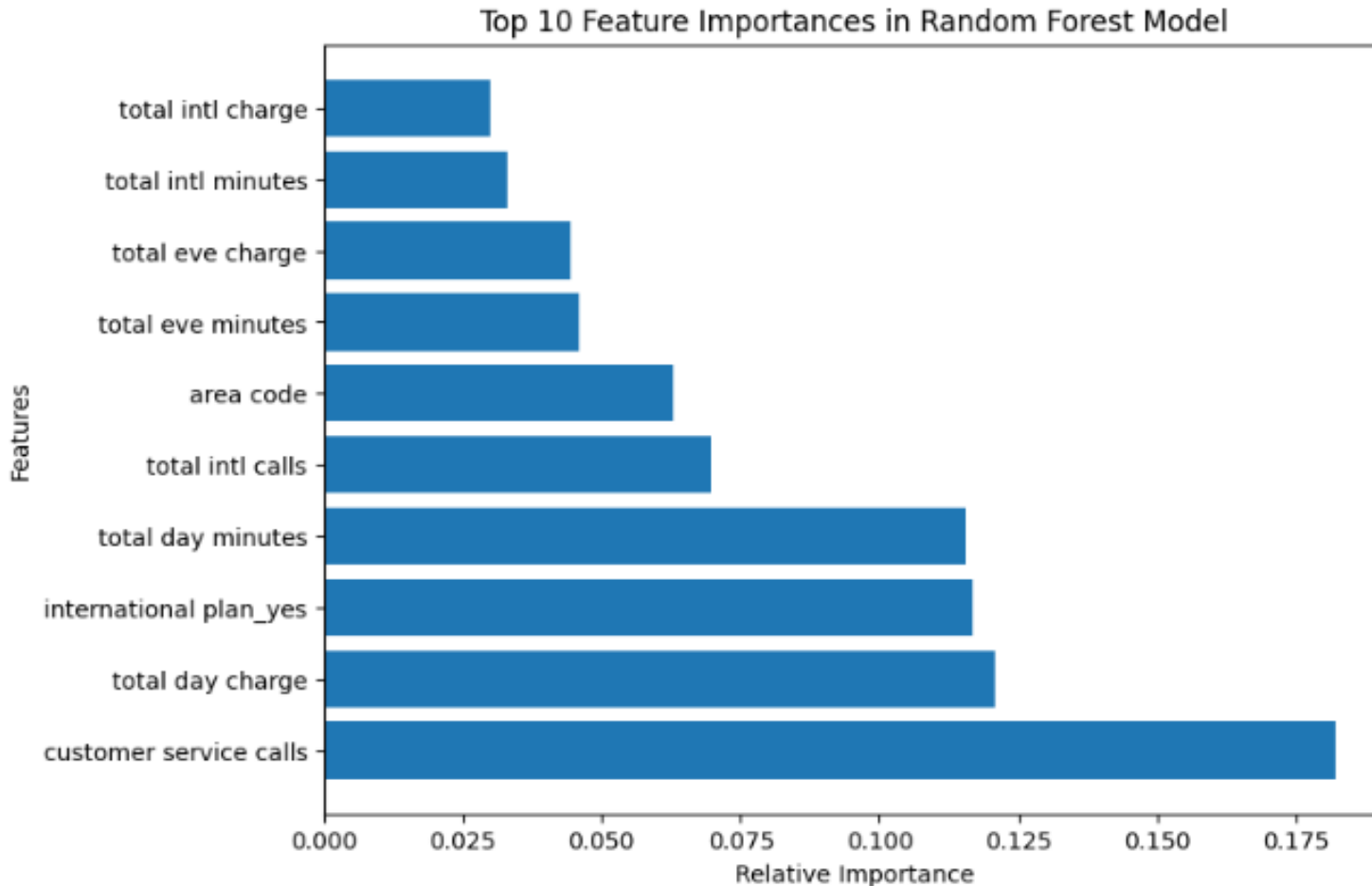
3

At **82%** recall score, performed marginally better in identifying churned customers

4

Achieved **95%** accuracy in correctly classifying the total number of customers

# Which features contribute most to customer churn?



- **Customer service calls** is the key factor contributing to high customer churn
- Other important features: **total day charge**, having an **international plan**, **total day minutes**, **total international calls**, and **area code**

# Conclusions

- Poor customer service was the leading cause of account terminations by customers
- Customers with an international plan churn more than those with no plan
- Customers with higher charges on day calls, especially those who have subscribed to more day minutes, are significantly more likely to churn
- The significantly higher charges observed for total daytime calls and minutes, in comparison to evening and nighttime calls and minutes, were a contributing factor to customer churn
- There is a lack of proportionality between the total number of international calls made and the corresponding charges, with higher charges even for fewer calls made

# Recommendations



**Prioritize customer service** initiatives to strengthen customer loyalty and satisfaction



**Review cost** of daytime calls and minutes charges



**Review pricing** plans to align with the market and retain price-sensitive customers

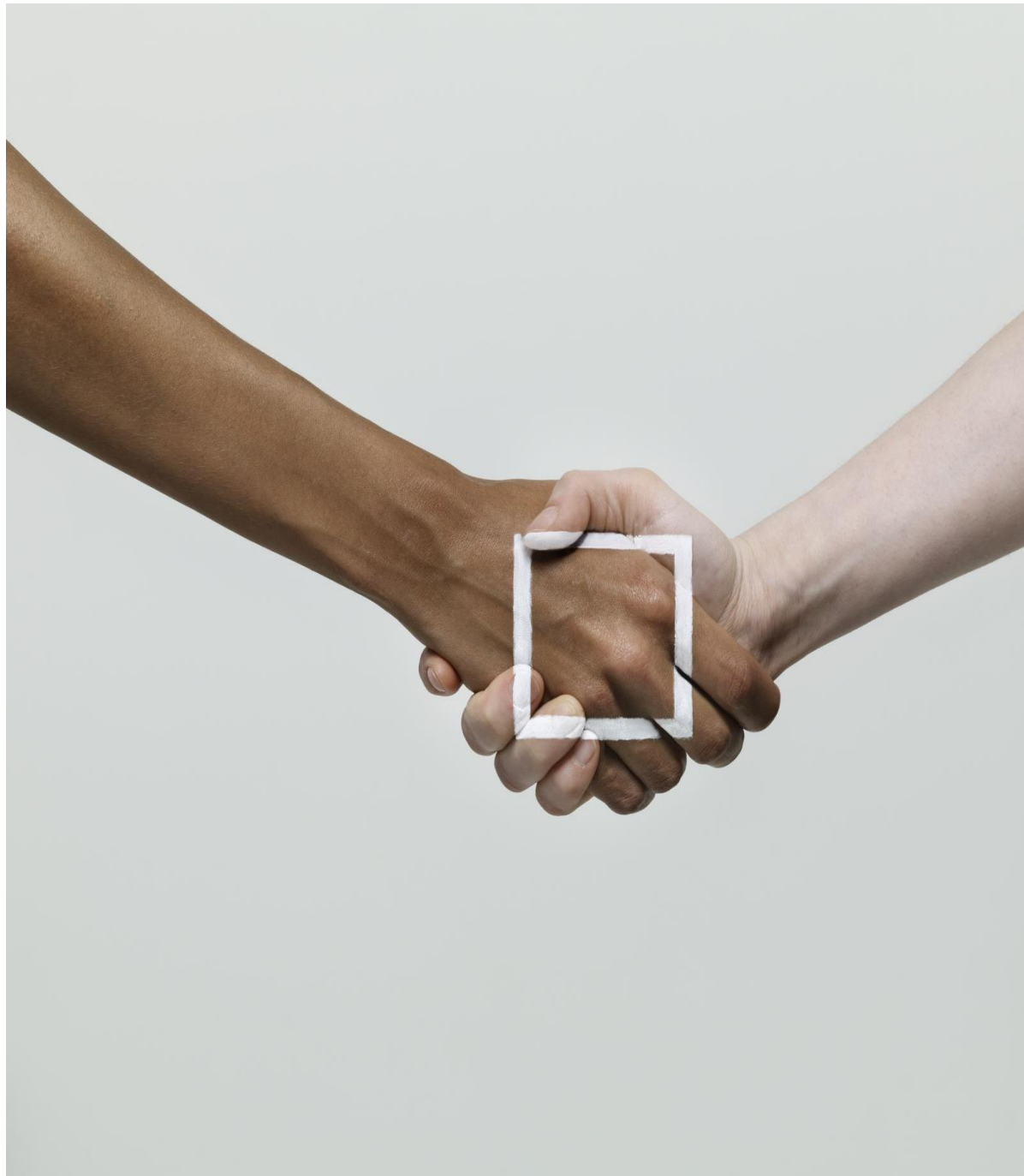


**Tailor personalized data and voice plan** products for international customers based on their unique needs



Regularly **gather customer feedback** to understand pain points and areas of improvement





# THANK YOU!

GitHub:

<https://github.com/Jay-Khainza>

<https://github.com/okwarojona>

<https://github.com/Kenkimmax>

<https://github.com/NazraJN>

<https://github.com/4kipkorir>